Direct Sum Theorem for Bounded Round Quantum Communication Complexity

Dave Touchette [1]

Abstract

We prove a direct sum theorem for bounded round entanglement-assisted quantum communication complexity. To do so, we use the fully quantum definition for information cost and complexity that we recently introduced, and use both the fact that information is a lower bound on communication, and the fact that a direct sum property holds for quantum information complexity. We then give a protocol for compressing a single copy of a protocol down to its quantum information cost, up to terms depending on the number of rounds and the allowed increase in error. Two important tools to derive this protocol are a smooth conditional min-entropy bound for a one-shot quantum state redistribution protocol, and the quantum substate theorem of Jain, Radhakrishnan and Sen (FOCS'02) to transform this bound into a von Neumann conditional entropy bound. This result further establishes the newly introduced notions of quantum information cost and complexity as the correct quantum generalisations of the classical ones in the standard communication complexity setting. Finding such a quantum generalisation of information complexity was one of the open problem recently raised by Braverman (STOC'12).

1 Introduction

We present the first general direct sum theorem for quantum communication complexity that holds for more than a single round of communication. A direct sum theorem states that to compute n tasks simultaneously requires as much resources as the amount of the given resource required for computing them separately. By a general direct sum theorem, we mean a direct sum theorem that holds for arbitrary relations on arbitrary inputs. The direct sum question, and the related direct product question, are of central importance in the different models of communication complexity. They have been the subject of a lot of attention in recent years. Many results were obtained for different models of classical communication complexity (see e.g. Refs [4, 16, 29, 32, 17, 14] and references therein). Progress for quantum communication complexity has been slower, with most results focusing on a single round of communication [35, 5, 28]. Some notable exceptions for the multi-round case are the work of Klauck, Špalek and de Wolf [39] in which they derive a direct product theorem for disjointness, and the works of Shaltiel [45], Lee, Shraibman and Špalek [40], and Sherstov [46] deriving direct product theorems for functions for which the discrepancy or generalized discrepancy method is tight. Even for a single round of communication, a general direct sum theorem was only proved earlier this year, using techniques much different from ours [2]. Previous to that work, techniques were restricted to proving results for the restricted case of product inputs. As a corollary of our results, we also obtain slightly improved parameters for the direct sum theorem of Ref. [2], for the single round case. The main tools that we use

[1] touchette.dave@gmail.com, Laboratoire d'informatique théorique et quantique, Département d'informatique et de recherche opérationnelle, Université de Montréal, Québec, Canada.

are new fully quantum notions of quantum information cost and complexity that we recently introduced [49], and a new single-shot protocol compression technique that we develop.

Quantum Information Complexity The classical notion of information cost was introduced by Chakrabarti, Shi, Wirth and Yao [18], who used it to derive a direct sum result for the simultaneous message passing model. The notion they introduced is similar to what is known today as the external information cost. A notion similar to what is now known as the internal information cost was later introduced by Bar-Yossef, Jayram, Kumar and Sivakumar [3] to use a direct sum property for composite problems that decompose into simpler ones, like the disjointness function in term of the AND function. The modern notions of external and internal information cost were formally introduced by Barak, Braverman, Chen and Rao [4], in which they prove general direct sum theorems for randomized communication complexity. For input random variables X and Y of Alice and Bob, respectively, shared randomness R, private randomness R_A, R_B available to Alice and Bob, respectively, and protocol transcript $\Pi(X, Y, R, R_A, R_B)$, the internal information cost is defined as $IC_{int}(\Pi, \mu) = I(X; \Pi|YR) + I(Y; \Pi|XR)$, and the external one as $IC_{ext}(\Pi, \mu) = I(XY; \Pi|R)$. Note that we have used Π to represent both the protocol and the protocol transcript, while μ is the prior distribution on the inputs X, Y. The interpretation of internal information cost is usually as the amount of information about Alice's input leaked to Bob plus the amount of information about Bob's input leaked to Alice, while for the external information it is as the amount of information about the joint input of Alice and Bob leaked to an external observer. Subsequent work by Braverman and Rao [16] provided an operational interpretation of internal information complexity as the amortized distributional communication complexity, i.e. the communication complexity per copy for computing n copies of a task in parallel, in the asymptotic limit of large n. They also provide a general direct sum theorem for bounded round communication complexity. Braverman [13] provides a similar operational interpretation of a prior-free version of information complexity as the amortized randomized communication complexity. He also list several interesting open questions related to information complexity, one of which is to develop a quantum analog of information complexity. He also asks whether the inherent reversibility of quantum computing, among other properties of quantum information, will impose a limit on the potential applications of such a quantity. Note that our results finally settle this: a notion of quantum information complexity with a similar operational interpretation and similar potential for applications as the classical one can indeed be defined.

In the quantum setting, many difficulties are immediately apparent in trying to generalize the classical definition. Firstly, by the no-cloning theorem [24, 53], there is no direct analogue for quantum communication of the notion of a transcript, available to all parties and containing all previous messages. In the entanglement assisted model, we can replace quantum communication by twice as much classical communication, by using teleportation [6]. However, if we consider the transcript obtained by replacing quantum communication by classical communication in this way, this transcript will be completely uncorrelated to the corresponding quantum messages and to the inputs. Indeed, the classical messages sent in the teleportation protocol are uniformly random, unless we take the remaining part of the EPR pair into account. A possible way around this might be to try to adapt the classical

definition by measuring the correlations between the inputs and the whole state, after reception of each message, of the receiving party. We can then even sum over the information contained in all messages. This yield a sensible notion of quantum information cost which is partly classical, and a similar quantity was used by Jain, Radhakrishnan and Sen to obtain a beautiful proof of a lower bound on the bounded round quantum communication complexity of the disjointness function [34]. A further variation on this was used by Jain and Nayak to obtain a lower bound for a variant of the Index function [31]. Work on direct sum results for a single round of communication also consider related notions [35, 28, 2]. However, these partly classical notions of quantum information cost all suffer from the drawback that they are only a lower bound on the communication cost once they have been divided by the number of messages. Then, the corresponding notion of quantum information complexity does not have the clear operational interpretation of classical information complexity as the amortized communication complexity, and is probably restricted to applications in bounded round scenarios.

We use the new, fully quantum notions of quantum information cost and complexity recently introduced in Ref. [49]. These are the first fully quantum definitions for such quantities. In particular, the notion of cost applies to arbitrary bipartite quantum protocols that are run on arbitrary bipartite quantum inputs, and the notion of complexity applies to arbitrary quantum tasks on arbitrary quantum input. Of particular interest in the setting of quantum communication complexity that we focus on in this work is the case of quantum protocols implementing classical tasks, e.g. evaluating arbitrary bipartite classical functions or relations on arbitrary bipartite input distributions below a specified error bound. However, the notion might also find applications for fully quantum tasks, for example quantum correlation complexity [37, 38], remote state preparation [27], or interactive variants of state redistribution [41, 23, 54] and its special cases of state merging [25, 26, 10, 1, 11], state splitting [1, 11], and source coding [44]. Note that the proof of our direct sum results can also be extended to such quantum tasks. To arrive at such a definition, we proposed a new interpretation of the classical internal information cost. Indeed, if we view each message generation in a protocol as a channel, then the information cost can be seen to be equal to the sum of the asymptotic costs of simulating many copies of each such channel with side information at the receiver and feedback to the sender [41], a task related to the reverse Shannon theorem [8, 52, 1, 7, 11]. Using known bounds for this task [41], this yield a strengthening of the classical amortized communication result for bounded round complexity [16, 13]. In the fully quantum setting, channel simulation, with side information at the receiver and with environment given as feedback to the sender, is equivalent to the state redistribution task. This insight led to the new, fully quantum definitions of information cost and complexity, and the link between state redistribution and one-shot protocol compression is then apparent. These new definitions are the firsts to satisfy all of the properties that we stated as desirable for these quantum notions. In particular, we proved the following properties in Ref. [49].

Theorem 1 ([49]) *The quantum information cost directly provides a lower bound on quantum communication cost for any protocol and input state, independent of the number of messages of the protocol (Lemma 1 in Ref. [49]).*

The corresponding quantum information complexity is exactly equal to the amortized

quantum communication complexity for any quantum task with fixed input state (Theorem 2 in Ref. [49]).

Quantum information complexity obeys an exact direct sum property (Corollary 3 in Ref. [49]).

For these last two results, they hold both for a fixed or unlimited number of messages.

Protocol Compression and Direct Sum To obtain our direct sum theorem, we first prove a protocol compression result stating that we can compress a single copy of a bounded round protocol proportionally to its information cost. A important ingredient in this proof is a single-message one-shot state redistribution protocol. A state redistribution protocol on input state ρ^{ABC}, with the A and C registers initially held by Alice, and the B register held by Bob, is a protocol that effectively transmits the C register to Bob while keeping the overall correlation with a purifying register R, up to some small error ε. We use a new achievability bound for a communication cost of $H^{\varepsilon}_{max}(C|B) - H^{\varepsilon}_{min}(C|BR) + O(\log(1/\varepsilon))$ [12]. The proof of this result appears in a joint work with Berta and Christandl [12]. Independently of our work, similar upper bounds on one-shot state redistribution have been obtained by Datta, Hsieh and Oppenheim [22].

We then use the substate theorem of Jain, Radhakrishnan and Sen [33, 36, 30] to transform this into a bound in terms of von Neumann conditional entropies, and what remains is a term proportional to the von Neumann conditional mutual information, as in asymptotic state redistribution. Our compression protocol applies this single message compression iteratively, and satisfies the following.

Theorem 2 *For each M-message protocol Π and input state ρ, there exists an M-message compression protocol Π' implementing Π on input ρ up to error $M\varepsilon$, and satisfying $QCC(\Pi') \in O((QIC(\Pi, \rho) + 1)/\varepsilon^2 + M/\varepsilon^2)$.*

By combining this protocol compression result with many properties of quantum information complexity in Theorem 1 above, we can obtain our main theorem, a direct sum theorem for bounded round quantum communication complexity that holds for all quantum tasks. Note that the theorem holds in the model in which we allow for arbitrary pre-shared entanglement. For concreteness, we state the result for classical relations.

Theorem 3 (main) *For any $\varepsilon_1, \cdots, \varepsilon_n, \varepsilon'' \in (0, 1/2)$, any relations $R_1, \cdots, R_n$ and any number of message M, $QCC^M(\otimes_i(R_i, \varepsilon_i)) \in \Omega(\sum_i((\frac{\varepsilon''}{M})^2 QCC^M(R_i, \varepsilon_i + \varepsilon'') - M))$. In particular, $QCC^M((R, \varepsilon)^{\otimes n}) \in \Omega(n((\frac{\varepsilon''}{M})^2 QCC^M(R, \varepsilon + \varepsilon'') - M))$.*

2 Preliminaries

2.1 Quantum Information Theory

We use the following notation for quantum theory; see [50, 51] for more details. We associate a quantum register A with a corresponding vector space, also denoted by A. We only consider finite-dimensional vector spaces. A state of quantum register A is represented by a density operator $\rho \in \mathcal{D}(A)$, with $\mathcal{D}(A)$ the set of all unit trace, positive semi-definite

linear operators mapping A into itself. We also will consider subnormalized states: let $\mathcal{D}_{\leq}(A)$ be the set of all positive semi-definite linear operators on A with trace at most one. More generally, we denote by $\mathcal{P}(A)$ the set of all positive semi-definite linear operators on A. We say that a state ρ is pure if it is a projection operator, i.e. $(\rho^{AR})^2 = \rho^{AR}$. For a pure state ρ, we often use the pure state formalism, and represent ρ by the vector $|\rho\rangle$ it projects upon, i.e. $\rho = |\rho\rangle\langle\rho|$. A quantum channel from quantum register A into quantum register B is represented by a super-operator $\mathcal{N}^{A\to B} \in \mathcal{C}(A, B)$, with $\mathcal{C}(A, B)$ the set of all completely positive, trace-preserving linear operators from $\mathcal{D}(A)$ into $D(B)$. If $A = B$, we might simply write $\mathcal{N}^A$, and when systems are clear from context, we might drop the superscripts. For channels $\mathcal{N}_1 \in \mathcal{C}(A, B), \mathcal{N}_2 \in \mathcal{C}(B, C)$ and state $\rho \in \mathcal{D}(A)$, we denote their composition as $\mathcal{N}_2 \circ \mathcal{N}_1 \in \mathcal{C}(A, C)$, with action $\mathcal{N}_2 \circ \mathcal{N}_1(\rho) = \mathcal{N}_2(\mathcal{N}_1(\rho))$. We might drop the $\circ$ if the composition is clear from context. For A and B isomorphic, we denote the identity mapping as $I^{A\to B}$, with some implicit choice for the change of basis. For $\mathcal{N}^{A_1\to B_1} \otimes I^{A_2\to B_2} \in C(A_1 \otimes A_2, B_1 \otimes B_2)$, we might abbreviate this as $\mathcal{N}$ and leave the identity channel implicit when the meaning is clear from context. An important subset of $\mathcal{C}(A, B)$ when A and B are isomorphic spaces is the set of unitary channels $\mathcal{U}(A, B)$, the set of all maps $U \in \mathcal{C}(A, B)$ with an adjoint map $U^\dagger \in \mathcal{C}(B, A)$ such that $U^\dagger \circ U = I^A$. Another important example of channel that we use is the partial trace $\mathrm{Tr}_B(\cdot) \in C(A \otimes B, A)$ which effectively gets rid of the B subsystem. Fixing a basis $\{|b\rangle\}$ for B, the action of Tr_B on any $\rho^{AB} \in D(A \otimes B)$ is $\mathrm{Tr}_B(\rho^{AB}) = \sum_b \langle b\,|\,\rho^{AB}\,|b\rangle$, and we write $\rho^A = \mathrm{Tr}_B(\rho^{AB})$. Note that the action of Tr_B is independent of the choice of basis chosen to represent it. We also denote $\mathrm{Tr}_{\neg A} = \mathrm{Tr}_B$ to express that we want to keep only the A register. Fixing a basis also allows us to talk about classical states and joint states: $\rho \in D(B)$ is classical (with respect to this basis) if it is diagonal in basis $\{|b\rangle\}$, i.e. $\rho = \sum_b p_B(b)|b\rangle\langle b|$ for some probability distribution p_B. More generally, subsystem B of ρ^{AB} is said to be classical if we can write $\rho^{AB} = \sum_b p_B(b)|b\rangle\langle b|^B \otimes \rho_b^A$ for some $\rho_b^A \in D(A)$. An important example of a channel mapping a quantum system to a classical one is the measurement channel Δ_B, defined as $\Delta_B(\rho) = \sum_b \langle b\,|\,\rho\,|b\rangle \cdot |b\rangle\langle b|^B$ for any $\rho \in D(B)$. Often, $A, B, C, \cdots$ will be used to discuss general systems, while $X, Y, Z, \cdots$ will be reserved for classical systems. For a state $\rho^A \in D(A)$, a purification is a pure state $\rho^{AR} \in \mathcal{D}(A \otimes R)$ satisfying $\mathrm{Tr}_R(\rho^{AR}) = \rho^A$. If R has dimension at least that of A, then such a purification always exists. For a given R, all purifications are equivalent up to unitaries. For a channel $\mathcal{N} \in C(A, B)$, a unitary extension is a unitary $U_{\mathcal{N}} \in U(A \otimes B', A' \otimes B)$ with $\mathrm{Tr}_{A'}(U_{\mathcal{N}}(\rho^A \otimes \sigma^{B'})) = \mathcal{N}(\rho^A)$ for some fixed $\sigma \in \mathcal{D}(B')$. It is sufficient to consider any fixed pure state σ. Such an extension always exists provided A' is of dimension at least $\dim(A)^2$ (note that we also must have $\dim(A) \cdot \dim(B') = \dim(A') \cdot \dim(B)$).

The notion of distance we use is the trace distance, defined for two states $\rho_1, \rho_2 \in D(A)$ as the sum of the absolute values of the eigenvalues of their difference:

$$\|\rho_1 - \rho_2\|_A = \mathrm{Tr}(|\rho_1 - \rho_2|).$$

It has an operational interpretation as four times the best bias possible in a state discrimination test between ρ_1 and ρ_2. The subscript tells on which subsystems the trace distance is evaluated, and remaining subsystems might need to be traced out. We use the following results about trace distance. For proofs of these and other standard results in quantum information theory that we use, see [51]. The trace distance is monotone under noisy channels:

for any $\rho_1, \rho_2 \in \mathcal{D}(A)$ and $\mathcal{N} \in C(A, B)$,

$$\|\mathcal{N}(\rho_1) - \mathcal{N}(\rho_2)\|_B \leq \|\rho_1 - \rho_2\|_A. \tag{2.1}$$

For unitaries, the equality becomes an identity, a property called unitary invariance of the trace distance. Hence, for any $\rho_1, \rho_2 \in D(A)$ and any $U \in \mathcal{U}(A, B)$, we have

$$\|U(\rho_1) - U(\rho_2)\|_B = \|\rho_1 - \rho_2\|_A. \tag{2.2}$$

Also, the trace distance cannot be increased by adjoining an uncorrelated system: for any $\rho_1, \rho_2 \in D(A), \sigma \in \mathcal{D}(B)$

$$\|\rho_1 \otimes \sigma - \rho_2 \otimes \sigma\|_{AB} = \|\rho_1 - \rho_2\|_A. \tag{2.3}$$

It follows that the trace distance obeys a property that we call joint linearity: for a classical system X and two states $\rho_1^{XA} = p_X(x)|x\rangle\langle x|^X \otimes \rho_{1,x}^A, \rho_2^{XA} = p_X(x)|x\rangle\langle x|^X \otimes \rho_{2,x}^A$,

$$\|\rho_1 - \rho_2\|_{XA} = \sum_x p_X(x)\|\rho_{1,x} - \rho_{2,x}\|_A. \tag{2.4}$$

The measure of information that we use is the von Neumann entropy, defined for any state $\rho \in D(A)$ as

$$H(A)_\rho = \mathrm{Tr}(\rho \log \rho),$$

in which we take the convention that $0 \log 0 = 0$, justified by a continuity argument. All logarithms are taken base 2. Note that H is invariant under unitaries applied on ρ. If the state to be evaluated is clear from context, we might drop the subscript. Conditional entropy for a state $\rho^{ABC} \in D(A \otimes B \otimes C)$ is then defined as

$$H(A|B)_{\rho^{AB}} = H(AB)_{\rho^{AB}} - H(B)_{\rho^B},$$

mutual information as

$$I(A;B)_{\rho^{AB}} = H(A)_{\rho^A} - H(A|B)_{\rho^{AB}},$$

and conditional mutual information as

$$I(A;B|C)_{\rho^{ABC}} = H(A|C)_{\rho^{AC}} - H(A|BC)_{\rho^{ABC}}.$$

Note that mutual information and conditional mutual information are symmetric in interchange of A, B. For any pure bipartite state $\rho^{AB} \in D(A \otimes B)$, the entropy on each subsystem is the same:

$$H(A) = H(B). \tag{2.5}$$

For a tripartite pure state $\rho^{ABC} \in \mathcal{D}(A \otimes B \otimes C)$, the conditional entropy satisfies a duality relation:

$$H(A|B) = -H(A|C). \tag{2.6}$$

For isomorphic A, A', a maximally entangled state $\psi \in \mathcal{D}(A \otimes A')$ is a pure state satisfying $H(A) = \log \dim(A) = \log \dim(A')$. For a system A of dimension $\dim(A)$ and any $\rho \in D(A \otimes B \otimes C)$, we have the bounds

$$0 \leq H(A) \leq \log \dim(A), \tag{2.7}$$
$$-H(A) \leq H(A|B) \leq H(A), \tag{2.8}$$
$$0 \leq I(A;B) \leq 2H(A), \tag{2.9}$$
$$0 \leq I(A;B|C) \leq 2H(A). \tag{2.10}$$

The conditional mutual information satisfy a chain rule: for any $\rho \in D(A \otimes B \otimes C \otimes D)$,

$$I(AB;C|D) = I(A;C|D) + I(B;C|AD). \tag{2.11}$$

For product states $\rho^{A_1B_1C_1A_2B_2C_2} = \rho_1^{A_1B_1C_1} \otimes \rho_2^{A_2B_2C_2}$, entropy is additive,

$$H(A_1A_2) = H(A_1) + H(A_2), \tag{2.12}$$

and so there is no conditional mutual information between product system,

$$I(A_1; A_2|B_1B_2) = 0, \tag{2.13}$$

and conditioning on a product system is useless,

$$I(A_1; B_1|C_1A_2) = I(A_1; B_1|C_1). \tag{2.14}$$

More generally,

$$I(A_1A_2; B_1B_2|C_1C_2) = I(A_1; B_1|C_1) + I(A_2; B_2|C_2). \tag{2.15}$$

Two important properties of the conditional mutual information are strong subadditivity and the data processing inequality: we consider an equivalent rewriting of strong subadditivity, which states that conditional mutual information is non-negative. For any $\rho \in \mathcal{D}(A \otimes B \otimes C)$ and $\mathcal{N} \in \mathcal{C}(B, B')$, with $\sigma = \mathcal{N}(\rho)$,

$$I(A;B|C)_\rho \geq 0, \tag{2.16}$$
$$I(A;B|C)_\rho \geq I(A;B'|C)_\sigma. \tag{2.17}$$

For classical systems, conditioning is equivalent to taking an average: for any $\rho^{ABCX} = \sum_x p_X(x)|x\rangle\langle x|^X \otimes \rho_x^{ABC}$, for a classical system X and some appropriate $\rho_x \in \mathcal{D}(A \otimes B \otimes C)$,

$$H(A|BX)_\rho = \sum_x p_X(x) H(A|B)_{\rho_x}, \tag{2.18}$$
$$I(A;B|CX)_\rho = \sum_x p_X(x) I(A;B|C)_{\rho_x}. \tag{2.19}$$

For one-shot state redistribution, we also make use of the following entropies and distance measure.

The relative entropy of $\rho \in \mathcal{D}_{\leq}(A)$ with respect to $\sigma \in \mathcal{P}(A)$ is defined as

$$D(\rho\|\sigma) = \mathrm{Tr}(\rho \log \rho) - \mathrm{Tr}(\rho \log \sigma). \tag{2.20}$$

Note that we can rewrite the conditional entropy of A given B for $\rho \in \mathcal{D}_{\leq}(A \otimes B)$ as

$$H(A|B)_\rho = -D(\rho^{AB}\|I^A \otimes \rho^B). \tag{2.21}$$

The max-relative entropy [20] of $\rho \in \mathcal{D}_{\leq}(A)$ with respect to $\sigma \in \mathcal{P}(A)$ is defined as

$$D_{max}(\rho\|\sigma) = \inf\{\lambda \in \mathbb{R} : 2^\lambda \sigma \geq \rho\}. \tag{2.22}$$

The conditional min-entropy [43] of A given B for $\rho \in \mathcal{D}_{\leq}(A \otimes B)$ is defined as

$$H_{min}(A|B)_\rho = -\inf_{\sigma^B \in \mathcal{D}(B)} D_{max}(\rho^{AB}\|I^A \otimes \sigma^B). \tag{2.23}$$

Based on the duality relation for conditional entropy, we define the conditional max-entropy of A given B for $\rho^{AB} \in \mathcal{D}_{\leq}(A \otimes B)$, purified by $\rho^{ABR} \in \mathcal{D}_{\leq}(A \otimes B \otimes R)$, as

$$H_{max}(A|B)_\rho = -H_{min}(A|R)_\rho. \tag{2.24}$$

Note that this is independent of the choice of the purification.

We also consider smooth version of these. The notion of distance used for smooth entropies is the purified distance [48], defined for $\rho, \sigma \in \mathcal{D}_{\leq}(A)$ as

$$P(\rho, \sigma) = \sqrt{1 - \bar{F}^2(\rho, \sigma)}. \tag{2.25}$$

Here, $\bar{F}(\rho, \sigma) = F(\rho, \sigma) + \sqrt{(1 - \mathrm{Tr}(\rho))(1 - \mathrm{Tr}(\sigma))}$, with $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_A$. Properties of the purified distance that we use are proved in Refs [48, 47]. Then, for $\rho \in \mathcal{D}_{\leq}(A)$,

$$\mathcal{B}^\varepsilon(\rho) = \{\bar{\rho} \in \mathcal{D}_{\leq}(A) : P(\rho, \bar{\rho}) \leq \varepsilon\}. \tag{2.26}$$

For $\varepsilon \geq 0$, the smooth conditional min-entropy of A given B for $\rho \in \mathcal{D}_{\leq}(A \otimes B)$ is defined as

$$H^\varepsilon_{min}(A|B)_\rho = \sup_{\bar{\rho}^{AB} \in \mathcal{B}^\varepsilon(\rho^{AB})} H_{min}(A|B)_{\bar{\rho}}, \tag{2.27}$$

and the smooth conditional max-entropy of A given B as

$$H^\varepsilon_{max}(A|B)_\rho = \inf_{\bar{\rho}^{AB} \in \mathcal{B}^\varepsilon(\rho^{AB})} H_{max}(A|B)_{\bar{\rho}}. \tag{2.28}$$

2.2 Quantum Communication Model

We focus in this note on the distributional quantum communication complexity of classical relations. The model for communication complexity that we consider is the following. For a given bipartite relation $T \subset X \times Y \times Z_A \times Z_B$ and input distribution μ on $X \times Y$, Alice and Bob are given input registers A_{in}, B_{in} containing their classical input $x \in X, y \in Y$

at the outset of the protocol, respectively and they output registers A_{out}, B_{out} containing their classical output $z_A \in Z_A, z_B \in Z_B$ at the end of the protocol, respectively, which should satisfy the relation R. We generally allow for some small error ε in the output, which will be formalized below. In this distributional communication complexity setting, the input is a classical state $\rho = \sum_{x\in X, y\in Y} \mu(x,y)|x\rangle\langle x|^{A_{in}} \otimes |y\rangle\langle y|^{B_{in}}$, similarly for the output $\Pi(\rho) = \sum_{z_A\in Z_A, z_B\in Z_B} p_{Z_AZ_B}(z_A, z_B)|z_A\rangle\langle z_A|^{A_{out}} \otimes |z_B\rangle\langle z_B|^{B_{out}}$ of the protocol Π implementing the relation, and the error parameter corresponds to the probability of failure $\sum_{x,y} \mu(x,y)[(x,y,\Pi(x,y)) \notin R] \leq \varepsilon$.

A protocol Π for implementing relation T on input $\rho^{A_{in}B_{in}}$ is defined by a sequence of unitaries $U_1, \cdots, U_{M+1}$ along with a pure state $\psi \in \mathcal{D}(T_A \otimes T_B)$ shared between Alice and Bob, for arbitrary finite dimensional registers T_A, T_B. For appropriate finite dimensional memory registers $A_1, A_3, \cdots A_{M-1}, A'$ held by Alice, $B_2, B_4, \cdots B_{M-2}, B'$ held by Bob, and communication registers $C_1, C_2, C_3, \cdots C_M$ exchanged by Alice and Bob, we have (see Figure 1 in Ref. [49]) $U_1 \in \mathcal{U}(A_{in} \otimes T_A, A_1 \otimes C_1), U_2 \in \mathcal{U}(B_{in} \otimes T_B \otimes C_1, B_2 \otimes C_2), U_3 \in \mathcal{U}(A_1 \otimes C_2, A_3 \otimes C_3), U_4 \in \mathcal{U}(B_2 \otimes C_3, B_4 \otimes C_4), \cdots, U_M \in \mathcal{U}(B_{M-2} \otimes C_{M-1}, B_{out} \otimes B' \otimes C_M), U_{M+1} \in \mathcal{U}(A_{M-1} \otimes C_M, A_{out} \otimes A')$. We slightly abuse notation and also write Π to denote the channel implemented by the protocol, i.e.

$$\Pi(\rho) = \mathrm{Tr}_{A'B'}(U_{M+1}U_M \cdots U_2U_1(\rho \otimes \psi)). \tag{2.29}$$

To formally define the error, we introduce a purification register R. For a classical input $\rho^{A_{in}B_{in}} = \sum_{x\in X, y\in Y} \mu(x,y)|x\rangle\langle x|^{A_{in}} \otimes |y\rangle\langle y|^{B_{in}}$ like we consider here, we can always take this purification to be of the form $|\rho\rangle^{A_{in}B_{in}R} = \sum_{x\in X, y\in Y} \sqrt{\mu(x,y)}\, |x\rangle^{A_{in}} |x\rangle^{A_{in}} |y\rangle^{B_{in}} |xy\rangle^{R_1} |xy\rangle^{R_2}$, for an appropriately chosen partition of R into R_1, R_2. Note that if we trace out the the R_2 register, then we are left with a classical state such that R_1 contains a copy of the joint input. Then we say that a protocol Π for implementing relation T on input $\rho^{A_{in}B_{in}}$, with purification $\rho^{A_{in}B_{in}R}$, has error $\varepsilon \in [0,1]$ if $P_e = \mathrm{Pr}_{\mu,\Pi}[\Pi(\rho^{A_{in}B_{in}R_1}) \notin T] \leq \varepsilon$. We denote the set of all such protocol as $\mathcal{T}(T, \mu, \varepsilon)$. If we want to restrict this set to bounded round protocols with M messages, we write $\mathcal{T}^M(T, \mu, \varepsilon)$. Note that, for simplicity, we only define protocols with an even number of messages; our results also hold without this restriction, though in the special case of one round protocols, we would rather consider relations with a single output to ensure that the quantum communication complexity is well-defined for all value of the error parameter. The introduction of the reference system R_1 serves to keep track of correlations between inputs and outputs in this distributional setting. Note that an alternate way to formulate the error criterion is by considering projectors $\Pi_g^T = \sum_{(x,y,z_A,z_B)\in T} |xyz_Az_B\rangle\langle xyz_Az_B|^{A_{out}B_{out}R_1}, \Pi_b^T = I^{A_{out}B_{out}R_1} - \Pi_g^T$, and then we can write $P_e = \mathrm{Tr}(\Pi_b^T \Pi(\rho^{A_{in}B_{in}R_1}))$.

Note that in the standard context of quantum communication complexity, our model would be akin to the model introduced by Cleve and Buhrman [19], with pre-shared entanglement (though the fact that we use quantum communication here instead of classical communication as in the original model could lead to an improvement up to a factor of two of the communication complexity, due to superdense coding [9], but no more, due to the teleportation protocol [6]), rather than to the model introduced by Yao [55], in which parties locally initialize their registers. This is the natural analogue of the framework for classical information complexity in which parties are allowed shared randomness for free, and this

seems to be necessary to obtain the operational interpretation of information complexity, classical and quantum, as the amortized communication complexity, and also for protocol compression. Known proofs of the additivity property rely heavily on the availability of shared resources to perform some kind of simulation. This is also true of many other interesting properties of information complexity. Also, the protocol compression methods that we propose rely on pre-shared entanglement to achieve good cost.

As was said before, our framework is the quantum generalization of the one for distributional information complexity. Let us formally define the different quantities that we work with.

Definition 1 *For a protocol* Π *as defined above, we define the* quantum communication cost *of* Π *as*

$$QCC(\Pi) = \sum_i \log \dim(C_i).$$

Note that we do not require that $\dim(C_i) = 2^k$ for some $k \in \mathbb{N}$, as is usually done. This will not affect our definition on information cost and complexity, but might affect the quantum communication complexity by at most a factor of two. The corresponding notion of quantum communication complexity of a relation is:

Definition 2 *For a relation* $T \subset X \times Y \times Z_A \times Z_B$*, an input distribution* μ *on* $X \times Y$ *and an error parameter* $\varepsilon \in [0, 1]$*, we define the* ε*-error* quantum communication complexity *of* T *on input* μ *as*

$$QCC(T, \mu, \varepsilon) = \min_{\Pi \in \mathcal{T}(T,\mu,\varepsilon)} QCC(\Pi).$$

Remark 1 *For any* $T, \mu, 0 \leq \varepsilon_1 \leq \varepsilon_2 \leq 1$*, the following holds:*

$$QCC(T, \mu, \varepsilon_2) \leq QCC(T, \mu, \varepsilon_1).$$

We have the following definition for bounded round quantum communication complexity, and a similar remark holds.

Definition 3 *For a relation* $T \subset X \times Y \times Z_A \times Z_B$*, an input distribution* μ *on* $X \times Y$ *and an error parameter* $\varepsilon \in [0, 1]$ *and a bound* $M \in \mathbb{N}$ *on the number of messages, we define the* M*-message,* ε*-error* quantum communication complexity *of* T *on input* μ *as*

$$QCC^M(T, \mu, \varepsilon) = \min_{\Pi \in \mathcal{T}^M(T,\mu,\varepsilon)} QCC(\Pi).$$

We are also interested in the quantum communication complexity of implementing multiple relations in parallel. A protocol Π_n is said to compute the n-fold product relation $T_1 \otimes T_2 \otimes \cdots \otimes T_n$ on input $\mu_1 \times \mu_2 \times \cdots \times \mu_n$, each with corresponding error ε_i, if for all $i \in [n]$,

$$P_e^i = \mathrm{Tr}(\Pi_b^{T_i} \, \mathrm{Tr}_{\neg A_{out}^i B_{out}^i R_1^i}(\Pi_n(\rho^{\otimes_k A_{in}^k B_{in}^k R_1^k}))) \leq \varepsilon_i \tag{2.30}$$

This error criterion corresponds to the one achieved when sequentially implementing the n relations T_i on respective input distribution μ_i and error ε_i, and, even for the case of $\varepsilon_i = \varepsilon$ for each i, this is weaker than demanding to simulate them with overall error $\varepsilon = \min_i \varepsilon_i$. Indeed, asking for overall error ε could be a much harder task. In particular, The direct product question, for $T_i = T$ for each i, asks if this overall error goes to 1 exponentially fast in n if we do not allow sufficiently more resources than for sequential implementation. Hence, if we want to obtain the intended operational interpretation, we have to settle for such a success parameter. We denote $\mathcal{T}_n(\otimes_i(T_i, \mu_i, \varepsilon_i))$ the set of all protocols achieving the above goal of having ε_i error for each relation, and can define the n-fold quantum communication complexity accordingly. We also specialize the definition to the special case where we are interested in implementing n times the same task, with $\mathcal{T}_n((T, \mu, \varepsilon)^{\otimes n})$ the corresponding set of protocols. We also consider bounded round variants; $\mathcal{T}_n^M(\otimes_i(T_i, \mu_i, \varepsilon_i)), \mathcal{T}_n^M((T, \mu, \varepsilon)^{\otimes n})$ for sets of protocols with at most M messages.

Definition 4 *For relations $T_i \subset X^i \times Y^i \times Z_A^i \times Z_B^i$, input distributions μ_i on $X^i \times Y^i$ and error parameters $\varepsilon_i \in [0, 1]$, we define the* n-fold quantum communication complexity *of $\otimes_i(T_i, \mu_i, \varepsilon_i)$ as*

$$QCC_n(\otimes_i(T_i, \mu_i, \varepsilon_i)) = \min_{\Pi_n \in \mathcal{T}_n(\otimes_i(T_i, \mu_i, \varepsilon_i))} QCC(\Pi_n).$$

Definition 5 *For a relation $T \subset X \times Y \times Z_A \times Z_B$, an input distribution μ on $X \times Y$ and an error parameter $\varepsilon \in [0, 1]$, we define the ε-error,* n-fold quantum communication complexity *of T on input μ as*

$$QCC_n((T, \mu, \varepsilon)^{\otimes n}) = \min_{\Pi_n \in \mathcal{T}_n((T, \mu, \varepsilon)^{\otimes n})} QCC(\Pi_n).$$

Note that for all n, $QCC_n(\otimes_i(T_i, \mu_i, \varepsilon_i)) \leq \sum_i QCC(T_i, \mu_i, \varepsilon_i)$, as is made clear by sequentially running the n protocols achieving the minimum in the definition of the quantum communication complexity. Restricting to performing the same task, we have $QCC_n((T, \mu, \varepsilon)^{\otimes n}) \leq nQCC(T, \mu, \varepsilon)$.

We have corresponding definitions for bounded round complexity.

Definition 6 *For relations $T_i \subset X^i \times Y^i \times Z_A^i \times Z_B^i$, input distributions μ_i on $X^i \times Y^i$, error parameters $\varepsilon_i \in [0, 1]$ and a bound $M \in \mathbb{N}$ on the number of messages, we define the M-message* n-fold quantum communication complexity *of $\otimes_i(T_i, \mu_i, \varepsilon_i)$ as*

$$QCC_n^M(\otimes_i(T_i, \mu_i, \varepsilon_i)) = \min_{\Pi_n \in \mathcal{T}_n^M(\otimes_i(T_i, \mu_i, \varepsilon_i))} QCC(\Pi_n).$$

Definition 7 *For a relation $T \subset X \times Y \times Z_A \times Z_B$, an input distribution μ on $X \times Y$, an error parameter $\varepsilon \in [0, 1]$ and a bound $M \in \mathbb{N}$ on the number of messages, we define the M-message, ε-error,* n-fold quantum communication complexity *of T on input μ as*

$$QCC_n^M((T, \mu, \varepsilon)^{\otimes n}) = \min_{\Pi_n \in \mathcal{T}_n^M((T, \mu, \varepsilon)^{\otimes n})} QCC(\Pi_n).$$

Let us now adapt the definition of quantum information cost and complexity from Ref. [49] to the setting we consider. The register R is the purification register, invariant throughout the protocol since we consider local isometric processing. Note that, as noted before when considering a R_1R_2 partition for R, for classical input distributions, the purification register can be thought of as containing a (quantum) copy of the classical input. The definition is however invariant under the choice of R and corresponding purification.

Definition 8 *For a protocol Π and an input distribution μ on $X \times Y$, we define the* quantum information cost *of Π on input μ as*

$$QIC(\Pi, \mu) = \sum_{i>0, odd} \frac{1}{2} I(C_i; R|B_{i-1}) + \sum_{i>0, even} \frac{1}{2} I(C_i; R|A_{i-1}),$$

in which we have labelled $B_0 = B_{in} \otimes T_B$.

Definition 9 *For a relation $T \subset X \times Y \times Z_A \times Z_B$, an input distribution μ on $X \times Y$ and an error parameter $\varepsilon \in [0, 1]$, we define the ε-error* quantum information complexity *of T on input μ as*

$$QIC(T, \mu, \varepsilon) = \inf_{\Pi \in \mathcal{T}(T, \mu, \varepsilon)} QIC(\Pi, \mu).$$

Note that taking an inf has already been proven to be necessary here in the analogous classical context [15]. The reason is that an infinite sequence of protocols, using more and more rounds, might indeed be necessary to asymptotically approach the quantum information complexity, with each message containing an infinitesimal amount of information.

Definition 10 *For a relation $T \subset X \times Y \times Z_A \times Z_B$, an input distribution μ on $X \times Y$, an error parameter $\varepsilon \in [0, 1]$ and a bound $M \in \mathbb{N}$ on the number of messages, we define the M-message, ε-error* quantum information complexity *of T on input μ as*

$$QIC^M(T, \mu, \varepsilon) = \inf_{\Pi \in \mathcal{T}^M(T, \mu, \varepsilon)} QIC(\Pi, \mu).$$

Taking inf here might also be necessary, but for a different reason: an infinite sequence of protocol might also be necessary, with larger and larger entanglement registers. This is somewhat related to the fact that no good bounds are known on the amount of entanglement required for the best protocols; see Ref. [42] and references therein for related discussions.

Note that the operational interpretation of quantum information complexity as the amortized quantum communication complexity extends to the distributional setting for classical relations that we study here.

3 Properties of Quantum Information Complexity

We first show that a direct sum property holds for quantum information complexity of classical relations, and that information lower bounds communication. In fact, these results, as well as most others in this work, can be extended to the quantum generalization of general

classical task, defined in Ref. [14] as any action on inputs that can be performed by a protocol. Of course, an error criterion appropriate for the quantum setting is then required; see next section for an example. We state the necessary lemmata, taken from Ref. [49], but do not give proofs since they follow exactly the same lines as the corresponding proofs given for implementing bipartite quantum channels in Ref. [49].

Lemma 1 ([49]) *For any protocol Π and input distribution μ, the following holds:*

$$0 \leq QIC(\Pi, \mu) \leq QCC(\Pi).$$

Lemma 2 ([49]) *For any two protocols Π^1 and Π^2 with M_1 and M_2 messages, respectively, there exists a M-message protocol Π_2, satisfying $\Pi_2 = \Pi^1 \otimes \Pi^2, M = \max(M_1, M_2)$, such that the following holds for any corresponding input states ρ^1, ρ^2:*

$$QIC(\Pi_2, \rho^1 \otimes \rho^2) = QIC(\Pi^1, \rho^1) + QIC(\Pi^2, \rho^2).$$

Lemma 3 ([49]) *For any M-message protocol Π_2 and any input states $\rho^1 \in \mathcal{D}(A^1_{in} \otimes B^1_{in}), \rho_2 \in \mathcal{D}(A^2_{in} \otimes B^2_{in})$, there exist M-message protocols Π^1, Π^2 satisfying $\Pi^1(\cdot) = \mathrm{Tr}_{A^2_{out}B^2_{out}} \circ \Pi_2(\cdot \otimes \rho^2), \Pi^2(\cdot) = \mathrm{Tr}_{A^1_{out}B^1_{out}} \circ \Pi_2(\rho^1 \otimes \cdot)$, and the following holds:*

$$QIC(\Pi^1, \rho^1) + QIC(\Pi^2, \rho^2) = QIC(\Pi_2, \rho^1 \otimes \rho^2).$$

Let us now set some notation. We say that a triple (T, μ, ε) is a task, corresponding to the implementation of the relation $T \subset X \times Y \times Z_A \times Z_B$ on input distribution μ on $X \times Y$ with error $\varepsilon \in [0, 1]$. We define a product task recursively, and with the following notation: a task is a product task, and if $T^1 = (T_1, \mu_1, \varepsilon_1) \otimes \cdots \otimes (\mathcal{T}_i, \mu_i, \varepsilon_i), T^2 = (T_{i+1}, \mu_{i+1}, \varepsilon_{i+1}) \otimes \cdots \otimes (T_n, \mu_n, \varepsilon_n)$ are two product tasks, then $T^1 \otimes T^2 = \bigotimes_{i \in [n]} (T_i, \mu_i, \varepsilon_i)$ is also a product task. We say that a protocol Π_n, with input space $A^1_{in} \otimes B^1_{in} \otimes \cdots \otimes A^n_{in} \otimes B^n_{in}$ and output space $A^1_{out} \otimes B^1_{out} \otimes \cdots \otimes A^n_{out} \otimes B^n_{out}$, succeeds at the product task $\bigotimes_i (T_i, \mu_i, \varepsilon_i)$ if it succeeds, for each i, at implementing relation T_i on input μ_i with error ε_i, and denote by $\mathcal{T}_\otimes(\bigotimes_i (T_i, \mu_i, \varepsilon_i))$ the set of all protocols achieving this. Once again, if we restrict this set to M-message protocols, we write $\mathcal{T}^M_\otimes(\bigotimes_i (T_i, \mu_i, \varepsilon_i))$. We then define the quantum information complexity of the product task $\bigotimes_i (T_i, \mu_i, \varepsilon_i)$ as

$$QIC_\otimes(\bigotimes_i (T_i, \mu_i, \varepsilon_i)) = \inf_{\Pi_n \in \mathcal{T}_\otimes(\bigotimes_i (T_i, \mu_i, \varepsilon_i))} QIC(\Pi_n, \mu_1 \otimes \cdots \otimes \mu_n). \tag{3.1}$$

For the bounded round variant, we have

$$QIC^M_\otimes(\bigotimes_i (T_i, \mu_i, \varepsilon_i)) = \inf_{\Pi_n \in \mathcal{T}^M_\otimes(\bigotimes_i (T_i, \mu_i, \varepsilon_i))} QIC(\Pi_n, \mu_1 \otimes \cdots \otimes \mu_n). \tag{3.2}$$

Corollary 1 *For any two product tasks T_1, T_2 and any bound $M \in \mathbb{N}$ on the number of messages,*

$$QIC_\otimes(T_1 \otimes T_2) = QIC_\otimes(T_1) + QIC_\otimes(T_2),$$
$$QIC^M_\otimes(T_1 \otimes T_2) = QIC^M_\otimes(T_1) + QIC^M_\otimes(T_2).$$

Corollary 2 *For relations $T_i \subset X^i \times Y^i \times Z_A^i \times Z_B^i$, input distributions μ_i on $X^i \times Y^i$, error parameters $\varepsilon_i \in [0,1]$ and any bound $M \in \mathbb{N}$ on the number of messages, the following holds:*

$$\begin{aligned} QIC_{\otimes}(\otimes_i(T_i, \mu_i, \varepsilon_i)) &= \sum_i QIC(T_i, \mu_i, \varepsilon_i), \\ QIC_{\otimes}^M(\otimes_i(T_i, \mu_i, \varepsilon_i)) &= \sum_i QIC^M(T_i, \mu_i, \varepsilon_i), \\ QIC_n((T, \mu, \varepsilon)^{\otimes n}) &= nQIC(T, \mu, \varepsilon), \\ QIC_n^M((T, \mu, \varepsilon)^{\otimes n}) &= nQIC^M(T, \mu, \varepsilon). \end{aligned}$$

Corollary 3 *For relations $T_i \subset X^i \times Y^i \times Z_A^i \times Z_B^i$, input distributions μ_i on $X^i \times Y^i$, error parameters $\varepsilon_i \in [0,1]$ and any bound $M \in \mathbb{N}$ on the number of messages, the following holds:*

$$\begin{aligned} QIC(\otimes_i(T_i, \mu_i, \varepsilon_i)) &\leq QCC(\otimes_i(T_i, \mu_i, \varepsilon_i)), \\ QIC^M(\otimes_i(T_i, \mu_i, \varepsilon_i)) &\leq QCC^M(\otimes_i(T_i, \mu_i, \varepsilon_i)), \\ QIC_n((T, \mu, \varepsilon)^{\otimes n}) &\leq QCC_n((T, \mu, \varepsilon)^{\otimes n}), \\ QIC_n^M((T, \mu, \varepsilon)^{\otimes n}) &\leq QCC_n^M((T, \mu, \varepsilon)^{\otimes n}). \end{aligned}$$

4 Single-Message Compression at Conditional Mutual Information

To be able to compress a protocol proportionally to its quantum information cost, we show how to compress a single message down to a communication cost proportional to its conditional mutual information, as in asymptotic state redistribution. Entanglement is deemed free for the compression.

We will make use of a one-shot state redistribution achievability bound that we derive, together with Berta and Christandl, in Ref. [12]. Note that, independently of our work, similar results were obtained by Datta, Hsieh and Oppenheim [22].

Let us first define the fully quantum task of quantum state redistribution [41, 23, 54] in a one-shot setting. We adopt the bipartite channel simulation viewpoint from Ref. [49], which is quite similar to the one taken here, but generalized to the fully quantum setting. We are interested in bipartite protocols that will implements, on a given input $\rho \in \mathcal{D}(A_{in} \otimes B_{in})$, the action of channel $\mathcal{N} \in \mathcal{C}(A_{in} \otimes B_{in}, A_{out} \otimes B_{out})$ on such an input, and then output $\mathcal{N}(\rho) \in \mathcal{D}(A_{out} \otimes \mathcal{B}_{out})$. Alice is given as input the A_{in} register, and outputs the A_{out} register, and similarly for Bob with the B_{in}, B_{out} registers. We generally allow for some small error in this channel simulation; this will be formalized below. In this section, we only consider one-message protocols, but this can be straightforwardly extended to multi-round protocols (as required in the next section) as we do in the semi-classical setting; also see Ref. [49]. Such a protocol Π is defined first by a pre-shared entangled state. Then, by a local operation of Alice, acting on her part of the input state and on her part of the entanglement, and generating her share of the output along with a quantum message to be communicated. Finally, by a local operation of Bob, acting on his part of the input, his part of the entanglement and also on the quantum message received from Alice, and

generating his share of the output. We denote by $QCC(\Pi)$ the size of the communication register, measured in qubits. We usually allow for arbitrary pre-shared entanglement, which does not count towards the communication cost. However, in this direct coding theorem, only EPR pairs are consumed and generated, and so it would also make sense to speak of the net entanglement cost of a protocol. The protocol Π can be seen to define a channel in $\mathcal{C}(A_{in} \otimes B_{in}, A_{out}, B_{out})$, which we also denote by Π. Then, we say that a protocol Π implements a channel $\mathcal{N}$ up to some error ε on some state ρ if, for a given purification $\rho^{A_{in}B_{in}R}$ of the input state, the following holds: $P(\mathcal{N}(\rho^{A_{in}B_{in}R}), \Pi(\rho^{A_{in}B_{in}R})) \leq \varepsilon$.

We can now state formally the definition of one-shot state redistribution.

Definition 11 *We say that a bipartite channel $\mathcal{N} \in \mathcal{C}(A_{in} \otimes B_{in}, A_{out} \otimes B_{out})$ implements* state redistribution *on input $\rho^{A_{in}B_{in}}$, with $A_{in} = A \otimes C, B_{in} = B, A_{out} = A, B_{out} = B \otimes C$, if it implements the identity channel on such a state and such a partition of the input-output registers, i.e. if it transfers the C part of ρ from Alice to Bob. For a protocol implementing this channel up to some error ε, we say that it is an ε-error state redistribution protocol.*

Our result in Ref. [12] is then the following.

Theorem 4 *For all $\varepsilon_1 \geq 0, \varepsilon_3 > 0, \rho \in \mathcal{D}(A \otimes B \otimes C)$ purified by $\rho^{ABCR} \in \mathcal{D}(A \otimes B \otimes C \otimes R)$, there exists a one-message, $(6\varepsilon_1 + 3\sqrt{3}\varepsilon_3)$-error one-shot state redistribution protocol Π with quantum communication satisfying*

$$QCC(\Pi) \leq \frac{1}{2} H_{max}^{\varepsilon_1}(C|B)_\rho - \frac{1}{2} H_{min}^{\varepsilon_1}(C|BR)_\rho + 4 \log \frac{1}{\varepsilon_3} + 2.$$

The above direct coding theorem is not tight in general for one-shot state redistribution (see discussion in Ref. [11]). It is however sufficient for our purpose. We now show how to apply the direct coding theorem along with the substate theorem of Jain, Radhakrishnan and Sen [33, 36, 30] to obtain a bound on one-shot state redistribution in terms of von Neumann conditional mutual information. This can then be applied iteratively in order to get bounded-round protocol compression proportional to the information cost. Let us first restate the substate theorem in the form that we will use, using the following definition for the smooth max-relative entropy of $\rho \in \mathcal{D}_{\leq}(A)$ with respect to $\sigma \in \mathcal{P}(A)$:

$$D_{max}^{\varepsilon}(\rho \| \sigma) = \inf_{\bar{\rho} \in B_\varepsilon(\rho)} D_{max}(\bar{\rho} \| \sigma). \tag{4.1}$$

Theorem 5 (Substate theorem [33, 36, 30]) *For $\rho \in \mathcal{D}(A), \sigma \in \mathcal{P}(A)$ and $\varepsilon \in (0, 1)$,*

$$D_{max}^{\varepsilon}(\rho \| \sigma) \leq (D(\rho \| \sigma) + 1)/\varepsilon^2 + \log(1/(1 - \varepsilon^2)).$$

Note that the statement is trivial when $\rho^0 \not\leq \sigma^0$. Also note that the square factor here is due to a difference in the distance function used (we use the purified distance) compared to the one of [30]. The smoothing parameter then changes accordingly. Smoothing is also done over a larger set here, since we allow for subnormalized states, and by consequence it is possible that smooth max-relative entropy is slightly smaller according to our definition

then to the one of Ref. [30]. This is however not an issue, since the smoothing is in the correct direction for the inequality in the substate theorem.

This lead to a lower bound on the conditional min-entropy, or equivalently, by the duality relations, to an upper bound on the conditional max-entropy, in terms of the conditional von Neumann entropy for a normalized state ρ:

Lemma 4 *For $\rho \in \mathcal{D}(A \otimes B)$ and $\varepsilon \in (0,1)$,*

$$H_{min}^{\varepsilon}(A|B)_\rho \geq (H(A|B)_\rho - 1)/\varepsilon^2 - \log(1/(1-\varepsilon^2)).$$

Proof.

$$\begin{aligned}
H_{min}^{\varepsilon}(A|B)_\rho &= \sup_{\bar{\rho}^{AB} \in B_\varepsilon(\rho^{AB})} H_{min}(A|B)_{\bar{\rho}} \\
&= \sup_{\bar{\rho}^{AB} \in B_\varepsilon(\rho^{AB})} (- \inf_{\sigma^B \in \mathcal{D}(B)} D_{max}(\bar{\rho}^{AB} \| I^A \otimes \sigma_B) \\
&\geq \sup_{\bar{\rho}^{AB} \in B_\varepsilon(\rho^{AB})} (-D_{max}(\bar{\rho}^{AB} \| I^A \otimes \rho^B)) \\
&= -D_{max}^{\varepsilon}(\rho^{AB} \| I^A \otimes \rho^B) \\
&\geq -(D(\rho^{AB} \| I^A \otimes \rho^B) + 1)/\varepsilon^2 - \log(1/(1-\varepsilon^2)) \\
&= (H(A|B)_\rho - 1)/\varepsilon^2 - \log(1/(1-\varepsilon^2)).
\end{aligned}$$

■

We get the following bound in terms of von Neumann conditional mutual information for one-shot state redistribution.

Lemma 5 *For all $\varepsilon_1 \in (0, 1/2)$, $\rho \in \mathcal{D}(A \otimes B \otimes C)$ purified by $\rho^{ABCR} \in \mathcal{D}(A \otimes B \otimes C \otimes R)$, there exists a one-message, $(7\varepsilon_1)$-error one-shot state redistribution protocol Π with quantum communication satisfying*

$$QCC(\Pi) \leq \frac{1}{2\varepsilon_1^2} I(C;R|B)_\rho + 2/\varepsilon_1^2 + 15.$$

Proof.

We take $\varepsilon_1 > 0$ and $\varepsilon_3 = \varepsilon_1/(3\sqrt{3})$ in Theorem 4. Note that for $x \geq 4, 2\log x \leq x$, and for $y \in (1,2), \log y \leq 1$. Then

$$\begin{aligned}
QCC(\Pi) \leq & \frac{1}{2} H_{max}^{\varepsilon_1}(C|B)_\rho - \frac{1}{2} H_{min}^{\varepsilon_1}(C|BR)_\rho + 4\log\frac{1}{\varepsilon_3} + 2 \\
= & -\frac{1}{2} H_{min}^{\varepsilon_1}(C|AR)_\rho - \frac{1}{2} H_{min}^{\varepsilon_1}(C|BR)_\rho + 4\log\frac{3\sqrt{3}}{\varepsilon_1} + 2 \\
\leq & \frac{1}{2}(-H(C|AR)_\rho + 1)/\varepsilon_1^2 + \frac{1}{2}\log(1/(1-\varepsilon_1^2) + \\
& \frac{1}{2}(-H(C|BR)_\rho + 1)/\varepsilon_1^2 + \frac{1}{2}\log(1/(1-\varepsilon_1^2) + 4\log\frac{1}{\varepsilon_1} + 12 + 2 \\
= & \frac{1}{2\varepsilon_1^2} I(C;R|B)_\rho + 1/\varepsilon_1^2 + 2\log\frac{1}{\varepsilon_1^2} + \log(1/(1-\varepsilon_1^2) + 14 \\
\leq & \frac{1}{2\varepsilon_1^2} I(C;R|B)_\rho + 2/\varepsilon_1^2 + 15.
\end{aligned}$$

■

5 Protocol Compression at Information Cost

We want to compress a single protocol to a quantum communication cost close to its quantum information cost. Entanglement is deemed free for the compression. To obtain such a one-shot compression protocol for bounded rounds protocol, we apply single message compression iteratively. We obtain the following result.

Lemma 6 *For any $\varepsilon_1 \in (0, 1/2)$, any M-message protocol Π, any input state ρ, there exists a M-message compression protocol Π' that implements protocol Π on input ρ up to error $7M\varepsilon_1$ and satisfies*

$$QCC(\Pi') \leq \frac{1}{\varepsilon_1^2} QIC(\Pi, \rho) + M(2/\varepsilon_1^2 + 16).$$

Proof. Define $\varepsilon' = 7\varepsilon_1$ and $t = 2/\varepsilon_1^2 + 15$. Given any M-message protocol Π and any state $\rho^{A_{in}B_{in}R}$, let

$$\rho_1^{A_1C_1B_0R} = U_1(\rho \otimes \psi), \rho_2^{A_1C_2B_2R} = U_2(\rho_1), \cdots, \rho_M^{A_{M-1}C_MB_MR} = U_M(\rho_{M-1})$$

in which we label $B_0 = B_{in} \otimes T_B, B_M = B_{out} \otimes B'$. Then, take $Q_i = \frac{1}{2\varepsilon_1^2} I(C_i; R|B_{i-1}) + t$ for i odd (we do not worry here about reusing the possibly generated entanglement, and simply discard it in the encoding and decoding maps to be defined below), $Q_i = \frac{1}{2\varepsilon_1^2} I(C_i; R|A_{i-1}) + t$ for i even. Then by Lemma 5 we have encoding and decoding maps E_i, D_i, along with corresponding entanglement $\psi_i \in D(T_A^i \otimes T_B^i)$ and communication register $\hat{C}^i$ of size $\dim \hat{C}^i = 2^{\lceil Q_i \rceil}$, with each satisfying

$$\|D_i \circ E_i(\rho_i \otimes \psi_i) - \rho_i\|_{A_iC_iB_{i-1}R} \leq \varepsilon' \tag{5.1}$$

for odd i, or

$$\|D_i \circ E_i(\rho_i \otimes \psi_i) - \rho_i\|_{A_{i-1}C_iB_iR} \leq \varepsilon' \tag{5.2}$$

for even i. Let $\hat{E}_i, \hat{D}_i$ be unitary extensions of E_i, D_i, respectively, requiring ancillary states $\sigma_i^E \in \mathcal{D}(E_i^{in}), \sigma_i^D \in \mathcal{D}(D_i^{in})$. We define the following protocol Π' starting from the protocol Π defined by $U_1, \cdots, U_{M+1}$ and ψ.

Protocol Π' on input σ:
-Take entangled state $\hat{\psi} = \psi \otimes \psi_1 \otimes \sigma_1^E \otimes \sigma_1^D \otimes \cdots \otimes \psi_M \otimes \sigma_M^E \otimes \sigma_M^D$.
-Take unitaries $\hat{U}_1 = \hat{E}_i \circ U_1, \hat{U}_2 = \hat{E}_2 \circ U_2 \circ \hat{D}_1, \cdots, \hat{U_M} = \hat{E}_M \circ U_M \circ \hat{D}_{M-1}, \hat{U}_{M+1} = U_{M+1} \circ \hat{D}_M$
-Take as output the A_{out}, B_{out} registers.

Note that the communication cost of Π' satisfies

$$\begin{aligned}
QCC(\Pi') &= \sum_i \log \dim(\hat{C}^i) \\
&= \sum_i \lceil Q_i \rceil \\
&\leq \sum_{i>0,odd} \frac{1}{2\varepsilon_1^2} I(C_i; R|B_{i-1}) + \sum_{i>0,even} \frac{1}{2\varepsilon_1^2} I(C_i; R|A_{i-1}) + M(t+1) \\
&= \frac{1}{\varepsilon_1^2} QIC(\Pi, \rho) + M(t+1).
\end{aligned}$$

This is also a M-message protocol, so is left to bound the error to make sure that Π' implements Π on ρ up to error $M\varepsilon'$. We have

$$\begin{aligned}
\|\Pi'(\rho) - \Pi(\rho)\| &= \| \operatorname{Tr}_{\neg A_{out}B_{out}} U_{M+1}\hat{D}_M\hat{E}_M U_M \hat{D}_{M-1} \cdots \hat{E}_1 U_1(\rho \otimes \hat{\psi}) \\
&\qquad - \operatorname{Tr}_{\neg A_{out}B_{out}} U_{M+1}U_M \cdots U_1(\rho \otimes \psi)\| \\
&= \| \operatorname{Tr}_{\neg A_{out}B_{out}} U_{M+1}D_M E_M U_M D_{M-1} \cdots E_1 U_1(\rho \otimes \psi \otimes \psi_1 \otimes \cdots \otimes \psi_M) \\
&\qquad - \operatorname{Tr}_{\neg A_{out}B_{out}} U_{M+1}U_M \cdots U_1(\rho \otimes \psi)\| \\
&\leq \| \operatorname{Tr}_{(A')(B')} U_{M+1}D_M \cdots E_2 U_2 D_1 E_1(\rho_1 \otimes \psi_1 \otimes \cdots \otimes \psi_M) \\
&\qquad - \operatorname{Tr}_{(A')(B')} U_{M+1}D_M \cdots E_2 U_2(\rho_1 \otimes \psi_2 \otimes \cdots \otimes \psi_M)\| \\
&\quad + \| \operatorname{Tr}_{(A')(B')} U_{M+1}D_M E_M U_M D_{M-1} \cdots U_3 D_2 E_2(\rho_2 \otimes \psi_2 \otimes \cdots \otimes \psi_M) \\
&\quad - \operatorname{Tr}_{(A')(B')} U_{M+1}D_M E_M U_M D_{M-1} \cdots E_3 U_3(\rho_2 \otimes \psi_3 \otimes \cdots \otimes \psi_M)\| \\
&\qquad + \cdots \\
&\quad + \| \operatorname{Tr}_{(A')(B')} U_{M+1}D_M E_M U_M D_{M-1} E_{M-1}(\rho_{M-1} \otimes \psi_{M-1} \otimes \psi_M) \\
&\qquad - \operatorname{Tr}_{(A')(B')} U_{M+1}D_M E_M U_M(\rho_{M-1} \otimes \psi_M)\| \\
&\quad + \| \operatorname{Tr}_{(A')(B')} U_{M+1}D_M E_M(\rho_M \otimes \psi_M) \\
&\qquad - \operatorname{Tr}_{(A')(B')} U_{M+1}(\rho_M)\| \\
&\leq \|D_1 E_1(\rho_1 \otimes \psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_M) - (\rho_1 \otimes \psi_2 \otimes \cdots \otimes \psi_M)\| \\
&\quad + \|D_2 E_2(\rho_2 \otimes \psi_2 \otimes \psi_3 \otimes \cdots \otimes \psi_M) - (\rho_2 \otimes \psi_3 \otimes \cdots \otimes \psi_M)\| \\
&\qquad + \cdots \\
&\quad + \|D_{M-1} E_{M-1}(\rho_{M-1} \otimes \psi_{M-1} \otimes \psi_M) - (\rho_{M-1} \otimes \psi_M)\| \\
&\quad + \|D_M E_M(\rho_M \otimes \psi_M) - (\rho_M)\| \\
&\qquad \leq M\varepsilon'.
\end{aligned}$$

The first equality is by definition, the second one by tracing the registers E_i^{out}, D_i^{out} from the unitary extensions to the encoders and decoders in the first term, the first inequality is by the triangle inequality and by definition of the ρ_i's, the second inequality is due to the monotonicity of trace distance under noisy channels, and the next is by (5.1) and (5.2), along with the fact that appending uncorrelated systems does not change the trace distance.

■

6 Direct Sum Theorem

Let us now introduce tasks for worst-case input: the task (T, ε) is similar to the task (T, μ, ε), but instead of requiring average error ε with respect to the input distribution μ, we require that for all inputs $(x, y) \in X \times Y$, the error is bounded by ε, i.e. $\Pr_\Pi[(x, y, \Pi(x, y)) \notin T] \leq \varepsilon$ for each pair (x, y). We get the following direct-sum theorem for bounded rounds entanglement assisted quantum communication complexity.

Theorem 6 *For any $\varepsilon, \varepsilon_i \in (0, 1/2)$, any relations T_i and any number of message M,*

$$QCC^M(\otimes_i(T_i, \varepsilon_i)) \geq \sum_i \varepsilon^2(QCC^M(T_i, \varepsilon_i + 14M\varepsilon) - M(2/\varepsilon^2 + 16)),$$
$$QCC^M((T, \varepsilon)^{\otimes n}) \geq n\varepsilon^2(QCC^M(T, \varepsilon + 14M\varepsilon) - M(2/\varepsilon^2 + 16)).$$

Proof. The second assertion clearly follows from the first, so we focus on the first. Define $\varepsilon' = 7\varepsilon$ and $t = 2/\varepsilon^2 + 16$. By the facts that, for M-message complexities, a direct sum theorem holds for quantum information and quantum information lower bounds quantum communication, i.e. Corollaries 2 and 3, we get that for any input distributions μ_i, $QCC^M(\otimes_i(T_i, \mu_i, \varepsilon_i)) \geq \sum_i QIC^M(T_i, \mu_i, \varepsilon_i)$. For any i and $\delta > 0$, consider now a M-message protocol Π_i such that $QIC(\Pi_i, \mu_i) \leq QIC^M(T_i, \mu_i, \varepsilon_i) + \delta$, and use Lemma 6 to obtain a M-message protocol Π_i' satisfying $QCC(\Pi_i') \leq \frac{1}{\varepsilon^2} QIC(\Pi_i, \mu_i) + Mt \leq \frac{1}{\varepsilon^2}(QIC^M(T_i, \mu_i, \varepsilon_i) + \delta) + Mt$, and such that Π_i' simulates Π_i on μ_i up to error $M\varepsilon'$. Trace distance is upper bounded by twice the purified distance, so, with ρ_i representing μ_i, we get $\|\Pi_i'(\rho_i) - \Pi_i(\rho_i)\|_{A_{out}B_{out}R_1} \leq 2M\varepsilon'$, also using monotonicity of the trace distance. Then, Π_i' implements T_i on μ_i up to error $\varepsilon_i + 2M\varepsilon'$:

$$\begin{aligned}\operatorname{Tr}(\Pi_b^{T_i}\Pi_i'(\rho_i^{A_{in}B_{in}R_1})) &\leq \operatorname{Tr}(\Pi_b^{T_i}\Pi_i(\rho_i^{A_{in}B_{in}R_1})) \\ &\quad + \|\operatorname{Tr}(\Pi_b^{T_i}\Pi_i'(\rho_i^{A_{in}B_{in}R_1})) - \operatorname{Tr}(\Pi_b^{T_i}\Pi_i(\rho_i^{A_{in}B_{in}R_1}))\|_{A_{out}B_{out}R_1} \\ &\leq \varepsilon_i + 2M\varepsilon'.\end{aligned}$$

Then $QCC^M(T_i, \mu_i, \varepsilon_i + 2M\varepsilon') \leq QCC(\Pi_i') \leq \frac{1}{\varepsilon^2}(QIC^M(T_i, \mu_i, \varepsilon_i) + \delta) + Mt$, and putting it all together, we get $\sum_i \varepsilon^2(QCC^M(T_i, \mu_i, \varepsilon_i + 2M\varepsilon') - Mt) - \delta \leq QCC^M(\otimes_i(T_i, \mu_i, \varepsilon_i))$. This holds for all $\delta > 0$, so taking the limit $\delta \to 0$ gives us the bound we are looking for, but with prior distributions μ_i. Then, the worst-case complexity upper bounds the distributional complexity for any input distribution, so $QCC^M(\otimes_i(T_i, \varepsilon_i)) \geq QCC^M(\otimes_i(T_i, \mu_i, \varepsilon_i))$. To conclude the proof, we can then use Yao's min-max principle to obtain a bound on the worst case communication complexity for each task in the sum. ■

By taking $\varepsilon = \varepsilon''/14M$ in the above theorem, we get the following formulation.

Corollary 4 *For any $\varepsilon_i, \varepsilon'' \in (0, 1/2)$, any relation T_i and any number of message M,*

$$QCC^M(\otimes_i(T_i, \varepsilon_i)) \geq \sum_i ((\frac{\varepsilon''}{14M})^2 QCC^M(T_i, \varepsilon_i + \varepsilon'') - 2M - 1),$$
$$QCC^M((T, \varepsilon)^{\otimes n}) \geq n((\frac{\varepsilon''}{14M})^2 QCC^M(T, \varepsilon + \varepsilon'') - 2M - 1).$$

7 Conclusion: Discussion and Open Questions

We prove the first general direct sum theorem for quantum communication complexity that holds for multiple rounds of communication. This had been an open question since the first works on the direct sum question for quantum communication [35]. The approach we took was to exploit the link between a new, fully quantum notion of quantum information complexity that we recently introduced [49], and the task of quantum state redistribution [41, 23, 54]. Indeed, obtaining such a direct sum result by appealing to a one-shot state redistribution protocol was one of the question left open in Ref. [49]. Obtaining a direct sum result for bounded round entanglement-assisted communication complexity was also one of the open question in Ref. [2]. The proof of the achievability bound on one-shot quantum state redistribution appears in a joint work with Berta and Christandl [12]. Independently, a similar bound was also derived in Ref. [22]. To connect the one-shot bounds given in terms of min- and max-entropies with the von Neumann conditional mutual information appearing in the quantum information cost, we make use of the substate theorem of Jain, Radhakrishnan and Sen [33, 36, 30].

There is possibly still room for improvement in the dependence on the number of rounds for the direct sum theorem that we prove, but new techniques will probably be required in order to get substantial improvement over the parameters that we obtain. The fact that we are doing compression in a message-by-message fashion, with non-negligible error for each message compression, and at fixed length encoding, impose severe limitations on the direct sum results that we can obtain.

However, the applicability of this notion of quantum information complexity to such a general, multi-round direct sum theorem, holding for all relations, provides further evidence that it is the correct quantum generalisation of classical information complexity to consider in the standard communication complexity setting. Along with the operational interpretation as the amortized communication complexity, and also its potential application to help in finally settling the bounded round quantum communication complexity of the disjointness function (see Ref. [34], and Sections 7 and 8 in Ref. [49] for partial results on this problem), it now appears clear that this is indeed the case. Finding such a quantum generalisation was one of the open questions stated by Braverman in Ref. [13].

Other potential applications of this notion of quantum information complexity is in obtaining time-space trade-off for quantum streaming algorithms [39, 31], obtaining the exact, up to second order, communication complexity of some problems, like the result in the classical setting that was recently obtained for the disjointness function [15]. Also, it would be interesting to investigate the general direct sum question in an unlimited round setting, and to try to obtain general direct product theorems, for which it is still an open question whether such a theorem holds even for the simplest case of a single round of quantum communication.

References

[1] Anura Abeyesinghe, Igor Devetak, Patrick Hayden, and Andreas Winter. *The mother of all protocols: Restructuring quantum information's family tree.* Proceedings of the Royal

Society of London. Series A (2009): 2537-2563.

[2] Anurag Anshu, Rahul Jain, Priyanka Mukhopadhyay, Ala Shayeghi, and Penghui Yao. *A new operational interpretation of relative entropy and trace distance between quantum states.* arXiv:quant-ph/1404.1366.

[3] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. *An information statistics approach to data stream and communication complexity.* Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science (2002): 209-218.

[4] Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. *How to compress interactive communication.* Proceedings of the 42nd Annual ACM Symposium on Theory of Computing (2010): 67-76.

[5] Avraham Ben-Aroya, Oded Regev, and Ronald de Wolf. *A hypercontractive inequality for matrix-valued functions with applications to quantum computing and LDCs* Proceedings of the Forty-Ninth Annual IEEE Symposium on Foundations of Computer Science (2008): 477-486.

[6] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wooters. *Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels.* Physical Review Letters 70.13 (1993): 1895-1899.

[7] Charles H. Bennett, Igor Devetak, Aram W. Harrow, Peter W. Shor, and Andreas Winter. *Quantum Reverse Shannon Theorem.* arXiv:quant-ph/0912.5537.

[8] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. *Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem.* IEEE Transactions on Information Theory 48.10 (2002): 2637-2655.

[9] Charles H. Bennett and Stephen J. Wiesner. *Communication via one- and two-particle operators on Einstein–Podolsky–Rosen states.* Physical Review Letters 69.20 (1992): 2881-2884.

[10] Mario Berta. *Single-shot quantum state merging.* Master's thesis, ETH Zurich, 2008.

[11] Mario Berta, Matthias Christandl, and Renato Renner. *The Quantum Reverse Shannon Theorem based on One-Shot Information Theory.* Commun. Math. Phys. 306, 579 (2011).

[12] Mario Berta, Matthias Christandl, Dave Touchette. *Smooth Entropy Bounds on One-Shot Quantum State Redistribution.* To simultaneously appear on the arXiv as the current work. (2014).

[13] Mark Braverman. *Interactive information complexity.* Proceedings of the 44th Annual ACM Symposium on Theory of Computing (2012): 505-524.

[14] Mark Braverman. *Interactive information and coding theory.* A survey accompanying an invited lecture at ICM'14 in Seoul.

[15] Mark Braverman, Ankit Garg, Denis Pankratov, and Omri Weinstein. *From information to exact communication.* Proceedings of the 45th Annual ACM Symposium on Theory of Computing (2013): 151-160.

[16] Mark Braverman, and Anup Rao. *Information equals amortized communication.* Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science (2011): 748-757.

[17] Mark Braverman, Anup Rao, Omri Weinstein, Amir Yehudayoff. *Direct products in communication complexity.* Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (2013).

[18] A. Chakrabarti, Yaoyun Shi ; A. Wirth, Andrew C.-C. Yao. *Informational complexity and the direct sum problem for simultaneous message complexity.* Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science (2001): 270-278.

[19] Richard Cleve, and Harry Buhrman. *Substituting quantum entanglement for communication.* Physical Review A 56.2 (1997): 1201-1204.

[20] Nilanjana Datta. *Min- and Max- Relative Entropies and a New Entanglement Monotone.* IEEE Transactions on Information Theory, vol. 55 (2009): 2816-2826

[21] Nilanjana Datta, and Min-Hsiu Hsieh. *The apex of the family tree of protocols: optimal rates and resource inequalities.* New Journal of Physics 13(9) (2011): 093042.

[22] Nilanjana Datta, Min-Hsiu Hsieh, and Jonathan Oppenheim. *An upper bound on the second order asymptotic expansion for the quantum communication cost of state redistribution.* Stated as in preparation in [21]. To simultaneously appear on the arXiv as the current work.

[23] Igor Devetak, and Jon Yard. *Exact cost of redistributing multipartite quantum states.* Physical Review Letters 100.23 (2008): 230501.

[24] Dennis Dieks. *Communication by EPR devices.* Physics Letters A 92.6 (1982):271-272.

[25] Michal Horodecki, Jonathan Oppenheim, Andreas Winter. *Partial quantum information.* Nature 436 (2005):673-676.

[26] Michal Horodecki, Jonathan Oppenheim, Andreas Winter. *Quantum state merging and negative information.* Comm. Math. Phys. 269, 107 (2007).

[27] Rahul Jain. *Communication complexity of remote state preparation with entanglement.* Quantum Information and Computation 6 (4,5), 461-464 (2006).

[28] Rahul Jain, Hartmut Klauck. *New results in the simultaneous message passing model via information theoretic techniques.* The 24th IEEE Conference on Computational Complexity (2009): 369-378.

[29] Rahul Jain, Hartmut Klauck, Ashwin Nayak. *Direct product theorems for classical communication complexity via subdistribution bounds.* The 40th ACM Symposium on Theory of Computing (2008): 599-608.

[30] Rahul Jain, and Ashwin Nayak. *Short proofs of the quantum Substate Theorem.* IEEE Trans. Inf. Theory 58(6), 3664 - 3669 (2012).

[31] Rahul Jain, and Ashwin Nayak. *The space complexity of recognizing well-parenthesized expressions in the streaming model: the Index function revisited.* To appear in: IEEE Transactions on Information Theory.

[32] Rahul Jain, Attila Pereszlnyi, and Penghui Yao. *A direct product theorem for bounded-round public-coin randomized communication complexity.* Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (2012): 167-176.

[33] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. *Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states.* Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science (2002): 429-438.

[34] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. *A lower bound for bounded round quantum communication complexity of set disjointness.* Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (2003): 220-229.

[35] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. *Prior entanglement, message compression and privacy in quantum communication.* The 20th IEEE Conference on Computational Complexity (2005): 285-296.

[36] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. *A new information-theoretic property about quantum states with an application to privacy in quantum communication.* Journal of ACM 56(6), 33 (2009).

[37] Rahul Jain, Yaoyun Shi, Zhaohui Wei, and Shengyu Zhang. *Efficient protocols for generating bipartite classical distributions and quantum states.* IEEE Transctions on Information Theory, 59 (2013):51715178.

[38] Rahul Jain, Zhaohui Wei, Penghui Yao, and Shengyu Zhang. *Multipartite Quantum Correlation and Communication Complexities.* arXiv:quant-ph/1405.6015 (2014).

[39] Hartmut Klauck, Robert Špalek, and Ronald de Wolf. *Quantum and classical strong direct product theorems and optimal time-space tradeoff.* SIAM J. Comput., 36 (2007): 1472-1493

[40] Troy Lee, Adi Shraibman, and Robert Špalek. *A direct product theorem for discrepancy.* Proceedings of the Twenty-Third Annual IEEE Conference on Computational Complexity (2008): 71-80.

[41] Zhicheng Luo, and Igor Devetak. *Channel Simulation With Quantum Side Information.* IEEE Trans. Inf. Theory 55(3), 1331-1342 (2009).

[42] Laura Mančinska, Thomas Vidick. *Unbounded entanglement can be needed to achieve the optimal success probability.* arXiv:quant-ph/1402.4145.

[43] Renato Renner *Security of Quantum Key Distribution.* Ph.D. thesis, ETH Zurich.

[44] Benjamin Schumacher. *Quantum coding* Physical Review A, 51(4):2738-2747, 1995.

[45] Ronen Shaltiel. *Towards proving strong direct product theorems.* Computational Complexity 12 (2003): 1-22.

[46] Alexander A. Sherstov. *Strong direct product theorems for quantum communication and query complexity.* Proceedings of the 43rd Annual ACM Symposium on Theory of Computing (2011): 41-50.

[47] Marco Tomamichel. *A Framework for Non-Asymptotic Quantum Information Theory.* Ph.D. thesis, ETH Zurich.

[48] Marco Tomamichel, Roger Colbeck, Renato Renner. *Duality Between Smooth Min- and Max-Entropies* IEEE Transactions on Information Theory 56, 4674-4681 (2010).

[49] Dave Touchette. *Quantum Information Complexity and Amortized Communication.* arxiv: quant-ph/1404.3733

[50] John Watrous. *Theory of Quantum Information.* Lecture notes from Fall 2013, https://cs.uwaterloo.ca/ watrous/CS766/ (2013).

[51] Mark M. Wilde. *Quantum Information Theory.* Cambridge University Press (2013), Preliminary version available as: arXiv e-print quant-ph/1106.1445.

[52] Andreas Winter. *Compression of sources of probability distributions and density operators* arXiv: quant-ph/0208131 (2002).

[53] William K. Wootters, and Wojciech H. Zurek. *A single quantum cannot be cloned.* Nature 299.5886 (1982): 802-803.

[54] Jon T. Yard, and Igor Devetak. *Optimal Quantum Source Coding With Quantum Side Information at the Encoder and Decoder.* IEEE Transactions on Information Theory 55.11 (2009): 5339-5351.

[55] Andrew C.-C. Yao. *Quantum circuit complexity.* Proceedings of the 34th Annual IEEE Symposium on Foundations of Computer Science (1993): 352-361.